

Le parole della statistica

»» Ampiezza delle classi

Differenza tra il confine superiore e il confine inferiore (v. *Confini delle classi*) di una classe di modalità.

L'ampiezza detta anche modulo, può essere uguale o diversa per tutte le classi.

»» Box-Plot

Il box-plot, altrimenti detto *grafico a scatola*, è una tipologia di rappresentazione proposta dallo statistico americano J. W. Tukey.

»» Campionamento

Rilevazione statistica che consiste nell'*estrazione* di un insieme definito di *unità statistiche* componenti la *popolazione*. Il ricorso alla rilevazione per *campione* è frequente poiché presenta numerosi vantaggi: minor costo, maggiore rapidità nella rilevazione e nella elaborazione dei dati e, per talune indagini, maggiore accuratezza e controllo nelle risposte. Si può distinguere tra *campionamento probabilistico* e *non probabilistico*: il primo ha come elemento qualificante la casualità ottenibile tramite l'attribuzione di una probabilità positiva di essere estratta ad ogni unità della popolazione ed impiegando tecniche di selezione casuale. Fanno parte di questa categoria: il *campionamento casuale semplice*, il *campionamento stratificato*, il *campionamento a più stadi*, il *campionamento a grappoli*. Nel campionamento non probabilistico, invece, manca l'elemento della casualità. In questo caso si menzionano: il *campionamento a scelta ragionata* ed il *campionamento per quote*.

»» Campionamento a grappoli

Procedura di *campionamento* che prevede la disposizione delle *unità statistiche* in grappoli (gruppi) e la successiva *estrazione* casuale dai gruppi. Entrano a far parte del *campione* tutti gli elementi contenuti in un gruppo selezionato.

»» Campionamento a più stadi

Procedura di *campionamento* che presuppone l'individuazione di una struttura gerarchica della *popolazione*, in cui le unità finali sono incluse in insiemi di livello via via più elevato. Un esempio di tale procedura è costituito dalla rilevazione periodica dell'*ISTAT* sulle forze lavoro.

»» Campionamento a probabilità variabile

Procedura di *campionamento* nella quale la *probabilità* di *estrazione* dell'*unità statistica* varia in funzione di un *carattere* ritenuto fondamentale per l'indagine.

»» Campionamento a scelta ragionata

Procedura di *campionamento* non probabilistico per cui vengono selezionate quelle *unità statistiche* che, a giudizio del ricercatore, meglio rappresentano il fenomeno in esame.

»» Campionamento casuale semplice

Tipo di *campionamento* in cui le unità della *popolazione* hanno la stessa *probabilità* di entrare a far parte del campione. Questa procedura non richiede alcuna informazione a priori sulle caratteristiche della popolazione e si basa su una tecnica di selezione casuale.

»» Campionamento per quote

Procedura di *campionamento* che consiste nell'affidare al rilevatore il compito di selezionare le unità del *campione* nel rispetto di quote di *popolazione* prefissate che presentano determinate caratteristiche. Esso si adotta nel caso di indagini su una popolazione distribuita su un territorio molto vasto per la quale non si possiede una lista completa dei suoi componenti.

»» Campionamento stratificato

Procedura di *campionamento* che consiste nel suddividere la *popolazione* in un numero determinato di strati o classi, in modo tale che le unità appartenenti ad uno strato siano omogenee rispetto al *carattere* in esame, e nell'estrarre un campione casuale semplice di numerosità prefissata da ciascuno strato.

»» Campione

Sottoinsieme di una *popolazione* ottenuto per *estrazione* tramite tecniche definite dalla teoria dei campioni. L'impiego di campioni di *unità statistiche* si basa sull'approccio induttivo tipico dell'*inferenza statistica*.

»» Campo di variazione

Differenza tra il valore massimo ed il valore minimo delle *modalità* di una *variabile*. È un indice assoluto di variabilità che prende in considerazione solo la dispersione esistente tra i valori estremi della distribuzione. Risente di eventuali *valori anomali*.

»» Carattere

Elemento che consente di descrivere una *popolazione* (o un *campione*). Esso può essere qualitativo, in tal caso si definisce anche *mutabile*, o quantitativo e si manifesta nell'*unità statistica* mediante *attributi*, nel caso qualitativo, o *modalità* nel caso quantitativo. Ad esempio, il carattere «età in anni compiuti» presenta le modalità 0, 1, 2... ecc.; il carattere «sesso» presenta gli attributi: «maschio» e «femmina».

»» Cartogramma

Sistema di *rappresentazione grafica* di dati statistici in cui l'intensità o l'andamento di un fenomeno sono rappresentati per mezzo del variare della gradazione di tinte o disegni, o del tipo di disegno o tratteggio.

I cartogrammi sono però dei grafici meno precisi, nella rappresentazione dei fenomeni, di quelli cartesiani che hanno valore più rigoroso.

»» Censimento

Rilevazione statistica effettuata su tutte le unità del collettivo. Nella maggior parte delle nazioni, esso è ripetuto con cadenza decennale sulla popolazione, sulle aziende agricole, sulle imprese commerciali, sulle abitazioni, e così via.

»» Classe di modalità

Ciascuno degli intervalli di prefissata ampiezza in cui risulta suddiviso l'insieme delle modalità di un carattere quantitativo X . Ad ogni classe si fa corrispondere una frequenza che indica il numero di unità della popolazione che possiedono un valore del carattere compreso tra i suoi limiti.

Una classe di modalità può essere:

- aperta sia a sinistra che a destra, in tal caso i limiti inferiore e superiore sono esclusi dalla classe;
- aperta a sinistra e chiusa a destra, in tal caso il limite inferiore è escluso dalla classe;
- chiusa a sinistra e aperta a destra, in tal caso il limite superiore è escluso dalla classe;
- chiusa sia a sinistra che a destra, in tal caso entrambi i limiti sono inclusi nella classe.

Talvolta, quando l'insieme dei valori del carattere non è strettamente specificabile, si tende a non precisare il limite inferiore e il limite superiore, rispettivamente, della prima e dell'ultima classe, in altre parole, la prima classe è aperta a sinistra e l'ultima è aperta a destra.

»» Codifica

Operazione di trasformazione dei dati dalla forma originaria in cui sono stati rilevati ad una più agevole per lo svolgimento di analisi statistiche (numeri o sequenze alfanumeriche). La codifica è spesso necessaria per sottoporre i dati qualitativi ad analisi statistiche standard.

»» Collettivo: [altro modo per indicare popolazione (v.)]

»» Confini delle classi

Sono gli estremi superiore e inferiore delle classi di modalità.

L'estremo superiore di una classe si ottiene dalla semisomma del limite superiore della classe data e del limite inferiore della classe successiva.

L'estremo inferiore di una classe, analogamente, si ottiene dalla semisomma del limite inferiore della classe data e del limite superiore della classe precedente.

»» Coordinate

Numeri usati per individuare la posizione di un punto in un sistema di riferimento fissato, su una retta, nel piano o nello spazio. È pertanto possibile parlare di:

- *Coordinata di un punto su una retta*. Data una retta, stabilito su di essa un punto (chiamato *origine*), un segmento da utilizzare quale unità di misura e assunto come positivo uno dei due versi di percorrenza, ad ogni punto P della retta resterà associato un unico numero reale x . Tale numero si definisce *coordinata ascissa* o, più semplicemente, *ascissa*. Esso esprime la misura, rispetto all'unità scelta, del segmento i cui estremi sono l'origine (O) e il punto P .
- *Coordinate cartesiane nel piano*. Si fissi un punto O nel piano (origine) e una coppia di rette perpendicolari, denominate, rispettivamente, asse delle x o delle *ascisse* e asse delle y o delle *ordinate*, passanti per esso. Si stabilisce, poi, un'unità di misura. Ciascun punto P del piano delimitato dalle rette sarà definito da una coppia ordinata di numeri reali (x, y) , detti *coordinate cartesiane*, che esprimono, ognuno la dimensione della *proiezione* del punto P sull'asse, rispettivamente, delle ascisse e delle ordinate.

»» Le parole della statistica

»» Curva normale

Curva di frequenza unimodale e simmetrica, per la quale media aritmetica, mediana e moda coincidono. Assume la caratteristica forma a campana ed è nota anche come curva di Gauss o curva gaussiana.

Fu proposta da Gauss nel 1809 nella teoria degli errori, ma è anche attribuita a Laplace che ne definì le principali proprietà prima della trattazione più completa fatta da Gauss.

»» Database

Collezione strutturata di dati organizzata in modo da renderla facilmente accessibile ad una vasta gamma di programmi di applicazione.

»» Dati

Informazioni qualitative e/o quantitative strutturate e trattate al fine di consentire l'elaborazione statistica.

»» Decili

Trattasi di quantili di una variabile che dividono la sua distribuzione con osservazioni ordinate in senso non decrescente in dieci parti uguali. Ad esempio, il primo decile di un fenomeno X è quel valore tale che il 10% della popolazione assume valori inferiori ad esso.

»» Devianza

Misura della variabilità di una distribuzione. È data dalla somma degli scarti dalla media aritmetica, al quadrato.

»» Deviazione standard [v. → Scarto quadratico medio]

»» Diagramma

Rappresentazione grafica utilizzata per valutare visivamente un fenomeno statistico, facilitando il lavoro di ricerca e la lettura delle tabelle statistiche.

»» Diagramma a barre

Noto anche con la denominazione di ortogramma, è utilizzato per rappresentare graficamente caratteri spesso qualitativi; può essere di due tipi: a colonne (verticale) o a nastri (orizzontale). Il grafico consiste in una successione di colonne o nastri equidistanti, tanti quanti sono le modalità del carattere, la cui altezza (o lunghezza) è proporzionale alla frequenza dell'attributo corrispondente.

»» Diagramma a segmenti

Rappresentazione grafica particolarmente adatta per caratteri quantitativi discreti e anche qualitativi. Si avvale, in un sistema di riferimento cartesiano, di segmenti rettilinei e paralleli aventi lunghezza proporzionale alle intensità o frequenze del carattere investigato.

»» Diagramma cartesiano

Grafico i cui punti (x_i, y_i) , che possono essere collegati o meno, sono individuati attraverso coordinate cartesiane, che rappresentano le caratteristiche sulle variabili X e Y rilevate sul collettivo di riferimento.

»» Diagramma circolare

Rappresentazione grafica circolare adatta per i caratteri qualitativi che mostra le proporzioni delle parti rispetto all'insieme.

Si basa sulla corrispondenza tra intensità o frequenze e superfici di cerchi o di settori circolari. Areogrammi per settori circolari (o grafici a torta). L'area del cerchio esprime l'intensità o la frequenza totale (espressa in termini assoluti o percentuali), invece i settori circolari esprimono le intensità o le frequenze delle varie modalità del carattere esaminato; essi vengono individuati dividendo l'intero angolo giro in parti proporzionali alle intensità o frequenze da rappresentare.

»» Distribuzione

Organizzazione di dati in forma tabellare che ad ogni modalità di un dato carattere fa corrispondere la rispettiva frequenza (distribuzione di frequenza), oppure che ad un insieme di numeri reali fa corrispondere un insieme di eventi (distribuzione di probabilità).

»» Distribuzione unimodale

Dicesi di distribuzione di frequenza che presenta una sola moda. Negli altri casi, si parla di distribuzioni bimodali, plurimodali.

»» Elaborazione dei dati

Applicazione di programmi statistici specifici ad un insieme di dati per l'ottenimento di informazioni sintetiche.

»» EUROSTAT [UFFICIO STATISTICO DELLE COMUNITÀ EUROPEE]

Ente statistico dell'Unione europea creato nel 1953, raccoglie e uniforma i dati statistici provenienti dagli Stati membri e dai paesi che compongono il SEE (Spazio economico europeo, area economica integrata che, in base ad accordi tra UE ed EFTA, è operativa dal gennaio 1994). I dati provenienti dagli istituti nazionali di statistica inerenti a questioni economiche, politiche e sociali relative all'Unione europea vengono difatti messi insieme e poi coordinati sulla base delle caratteristiche peculiari di ciascun Paese.

»» Fonti

Qualsiasi tipo di documentazione che raccoglie dati riferiti ai fenomeni oggetto di studio. Possono essere di diverso tipo: orali, documentali, statistiche.

»» Fonti statistiche

Fonti che forniscono le informazioni sotto forma di dati statistici.

»» Frequenza assoluta

Designata anche semplicemente con il termine frequenza, indica il numero di unità statistiche che presentano la stessa modalità del carattere considerato. La indichiamo con n_i . Per esempio, poiché la popolazione femminile residente in Italia al 1° gennaio 2005 era pari a 30.085.571 unità, allora 30.085.571 è la frequenza assoluta del carattere «Numero di unità residenti in Italia al 1° gennaio 2005» relativo alla modalità $X =$ femmina.

»» Frequenza condizionata

In una tabella a doppia entrata, indica il numero di elementi che possiedono una determinata modalità di un carattere dato il valore assunto dall'altro carattere.

»» Frequenza cumulata

Rappresenta l'ammontare del carattere posseduto dalle prime i modalità, ordinate in senso non decrescente.

»» Frequenza relativa

Rapporto tra la frequenza assoluta ed il numero totale di unità statistiche del collettivo.

»» Gauss, Karl Friederich (Brunswick, 1777- Gottinga, 1855)

Matematico, probabilista, fisico e astronomo tedesco. È ritenuto uno dei più grandi matematici della storia ed ha contribuito per molti aspetti alla teoria della probabilità e alla statistica con una serie di importanti ed originali risultati.

In modo esplicito, sono attribuiti a Gauss la teoria degli errori accidentali, il metodo dei minimi quadrati, la definizione di distorsione e molte proprietà della variabile casuale normale, che porta il suo nome.

»» Ideogramma

Diagramma nel quale la rappresentazione avviene attraverso raffigurazioni di oggetti che ricordano il fenomeno da studiare; in pratica, con l'ideogramma si sostituisce all'unità di misura un'icona il più attinente possibile al fenomeno oggetto di analisi.

»» Indici di posizione

Sintetizzano la posizione di una distribuzione di frequenza mediante un valore reale «rappresentativo» della globalità del fenomeno e tale da riassumerne gli aspetti ritenuti più importanti. Tra i più importanti, citiamo la media aritmetica, la moda, la mediana, il valore centrale.

»» Indici di variabilità

Misurano la variabilità di una distribuzione di frequenza:

- rispetto ad un centro rappresentativo (*dispersione*); sono detti *scostamenti medi* e si ottengono determinando gli scarti tra le modalità del carattere e una sua *media*, e sintetizzandoli con la stessa o con altra *media*;
- tra le *unità statistiche* a due a due (*disuguaglianza*); sono detti *differenze medie* e si ottengono determinando le differenze in valore assoluto delle modalità del carattere prese a due a due, e sintetizzandole con opportuna *media*.

»» Indici statistici

Misure sintetiche introdotte per valutare aspetti complessi e globali di una distribuzione di un fenomeno X mediante un solo numero reale costruito in modo da disperdere al minimo le informazioni sui dati originari. In rapporto alle caratteristiche che si misurano, si parla di indici di posizione, indici di variabilità, indici di forma.

»» Intensità

Numero di unità statistiche esprimenti l'ammontare o la misura di un carattere. L'intensità è il risultato di una misurazione. Sono esempi di intensità il reddito di un individuo, il peso di una persona ecc.

»» ISTAT [ISTITUTO NAZIONALE DI STATISTICA]

Istituito nel 1926, è persona giuridica di diritto pubblico con ordinamento autonomo ed è sottoposto alla vigilanza della Presidenza del Consiglio dei Ministri e al controllo della Corte dei Conti. È l'organo statistico ufficiale dello Stato e provvede alla rilevazione, elaborazione e diffusione delle statistiche periodiche ed occasionali riguardanti diversi aspetti della vita del paese: demografico, sanitario, sociale, economico ecc.

»» Istogramma

Rappresentazione grafica di caratteri quantitativi continui le cui modalità sono costituite da classi di valori che possono essere di ampiezza uguale o meno. Nel primo caso, dall'asse orizzontale si sviluppano tanti rettangoli contigui, ciascuno avente base uguale all'ampiezza della classe e altezza uguale o proporzionale alla frequenza assoluta o relativa della classe corrispondente. Nel secondo caso, si hanno una serie di rettangoli contigui costruiti in modo che l'area di ogni rettangolo sia pari alla frequenza assoluta o relativa di ciascuna classe di modalità.

»» Media

Indice statistico di sintesi di una distribuzione compreso tra il minimo ed il massimo. Consente con un solo valore di sintetizzare l'ordine di grandezza di un dato fenomeno e confrontarlo con altri.

»» Media aritmetica

È l'indice statistico più utilizzato per la facilità di calcolo e per le proprietà di cui gode. Esprime la posizione globale di una distribuzione di frequenza. Date le modalità X_1, X_2, \dots, X_K che si presentano con frequenze assolute N_1, N_2, \dots, N_K , la media è rappresentata dal valore che si ottiene dividendo la loro somma per il loro numero N .

Un difetto della media aritmetica (rispetto alla mediana, per esempio) è la sua scarsissima resistenza ai valori eccezionali.

»» Media ponderata

Si parla di media ponderata quando si attribuisce diversa importanza alle modalità di una distribuzione assegnando un peso a ciascuna di esse.

»» Mediana

Indice di posizione. La mediana di una distribuzione di frequenza è quella modalità, in un insieme di osservazioni poste in ordine non decrescente, che divide la distribuzione in due parti (gruppi) di eguale ampiezza. Per cui, al primo gruppo apparterranno le osservazioni inferiori o uguali alla mediana; al secondo gruppo le osservazioni superiori o uguali alla stessa. Quando il numero delle osservazioni è dispari, la mediana sarà il valore che occupa la posizione centrale nella distribuzione; quando è pari, si calcola la semisomma dei due valori centrali.

Rispetto alla media aritmetica, la mediana è un indice più resistente perché non cambia se un valore eccezionale è presente nella distribuzione.

»» Moda

Indice di posizione; è la modalità del carattere cui corrisponde la massima frequenza. Ha significato di sintesi quando la distribuzione è unimodale.

»» Le parole della statistica

Per caratteri discreti la moda si identifica facilmente, scorrendo lungo la colonna delle frequenze.

Per caratteri continui se le classi di modalità hanno uguale ampiezza, la moda cade in quella con maggiore frequenza.

»» Modalità

Sono i diversi valori o qualità che può assumere un carattere sull'unità statistica. Quando le modalità sono di tipo qualitativo si parla anche di attributi.

»» Mutabili

Si parla di mutabili quando, nell'ambito di una rilevazione statistica, il carattere considerato presenta modalità qualitative. Ad esempio: sesso, professioni, colori ecc. sono mutabili. Le modalità che le mutabili assumono sono, spesso, dette attributi. Le mutabili si indicano anche come variabili qualitative.

»» Numerosità campionaria

Numero n di unità statistiche della popolazione costituenti il campione.

Talvolta, si distingue tra grandi e piccoli campioni a seconda che sia $n > 30$ oppure $n < 30$, rispettivamente.

»» Ortogramma

È sinonimo di diagramma a barre.

»» Percentile

Insieme di valori che dividono una distribuzione in cento parti uguali in modo che se x_1 è il primo percentile allora l'1% della popolazione possiede il carattere considerato X in misura inferiore o uguale a x_1 (v. Quantili).

»» Pictogramma

È sinonimo di ideogramma (v.)

»» Poligono di frequenza

Rappresentazione grafica di un carattere continuo che si avvale di un sistema di riferimento cartesiano (v. Diagramma cartesiano). Quando l'ampiezza delle classi diviene infinitesima, il poligono di frequenza si approssima a una curva regolare, la curva di frequenza.

»» Popolazione

Aggregato finito o infinito di unità (non necessariamente riferito a esseri umani) che forma l'oggetto dello studio statistico.

»» Quantili

Indici di posizione che dividono una distribuzione di frequenza in un certo numero di parti uguali, dopo aver ordinato i dati. Si hanno:

- i *terzili*, in numero di 2; il primo terzile (T_1) è quell'indice che lascia alla sua sinistra $1/3$ delle osservazioni, pari al 33,33%, e alla sua destra i rimanenti $2/3$, pari al 66,67%; il secondo terzile (T_2) lascia alla sua sinistra $2/3$ e alla sua destra $1/3$ delle osservazioni;

- i *quartili*, in numero di 3; il primo quartile (Q_1) lascia alla sua sinistra 1/4 delle osservazioni, pari al 25%, e alla sua destra i rimanenti 3/4, pari al 75%; ragionamento analogo vale per il secondo quartile (Q_2) che coincide con la *mediana* (v.) e per il terzo quartile (Q_3);
- i *decili*, in numero di 9;
- i *centili*, in numero di 99;

»» Quartili

Indici di posizione di una distribuzione statistica, sono particolari tipologie di quantili.

»» Quoziente

Indica il rapporto ovvero il confronto tra due grandezze omogenee, finalizzato alla stima della struttura e dell'andamento del fenomeno in esame (v. Rapporti statistici).

»» Range: [v. campo di variazione]

»» Rapporti statistici

Rapporti che consentono di confrontare intensità o frequenze di due fenomeni. I rapporti posti in essere tra le cifre rappresentanti i due fenomeni non sono mai fini a se stessi, ma rappresentano la traduzione matematica di rapporti logici che esistono tra i due fenomeni. Eliminano l'effetto dell'unità di misura, dando luogo a numeri puri che prescindono da essa.

»» Scarto

Differenza tra il valore assunto da una variabile ed un valore prefissato comunemente posto pari a un indice di posizione. Si parla, così di scarto dalla media aritmetica, scarto della mediana.

»» Scarto quadratico medio

Noto anche con la denominazione di deviazione standard, è un indice di variabilità indicato con σ ; è dato dalla radice quadrata della media aritmetica dei quadrati di tutte le differenze tra i valori della distribuzione e la media della distribuzione.

»» Scostamento

In statistica è sinonimo di scarto. La dizione viene anche utilizzata per indicare numerosi indici di variabilità.

»» Scostamento medio dalla media aritmetica

Indice di variabilità dato dalla media aritmetica dei valori assoluti degli scarti dalla media aritmetica.

È chiamato anche scarto semplice medio.

»» Scostamento medio dalla mediana

Indice di variabilità dato dalla media aritmetica dei valori assoluti degli scarti dalla mediana.

»» Serie cronologica [v. → Serie storica]

»» Serie storica

Una serie storica, x_t , è una successione di osservazioni ordinate logicamente secondo un indice temporale t che definisce l'ordinamento dei dati.

L'ordine degli elementi di una serie storica non è casuale, anzi i dati sono ordinati in modo naturale dal valore assunto dall'indice temporale. Fondamentale, nell'analisi delle serie storiche, è la scelta dei periodi di osservazione: tutti devono essere contraddistinti da omogeneità di caratteristiche.

La periodicità delle osservazioni è variabile; generalmente le serie storiche sono mensili, trimestrali o annuali. Nell'analisi economica, tuttavia, si incontrano spesso serie settimanali, giornaliere o orarie, o, al contrario, serie biennali, decennali.

»» Serie temporale [v. → Serie storica]

»» Serie territoriali

Distribuzioni statistiche nelle quali le modalità del carattere considerato sono ordinate secondo luoghi, territori, unità amministrative.

»» Simmetria

Si dice simmetrica una distribuzione rispetto al suo baricentro se le modalità che sono equidistanti dal valore centrale hanno la stessa frequenza. In una distribuzione simmetrica media aritmetica, mediana e moda coincidono.

»» Statistica

Scienza, strumentale ad altre, che elabora i metodi più idonei per analizzare fenomeni collettivi o di massa a fini descrittivi, esplicativi, previsionali ecc. Nella sua accezione più moderna, la statistica è la scienza delle decisioni da assumere in condizioni di incertezza e si articola nelle analisi esplorative dei dati, nello studio probabilistico e inferenziale dei risultati campionari, nella costruzione di un modello statistico.

»» Statistica descrittiva

Branca della statistica rivolta alla descrizione dei dati rilevati su uno o più fenomeni reali.

»» Statistica inferenziale

Branca della statistica rivolta all'induzione probabilistica circa la popolazione sulla base di misurazioni o osservazioni su un campione di individui prescelti con procedure casuali dalla popolazione. In termini generali, la statistica inferenziale è volta ad affrontare e risolvere il cosiddetto problema inverso in cui la struttura della popolazione è incognita e induttivamente si cerca di utilizzare un insieme limitato di informazioni fornite da un campione casuale estratto dalla popolazione.

»» Stereogramma

Rappresentazione grafica di due caratteri presentati in classi di modalità, costituita da un insieme di parallelepipedi contigui, in cui le dimensioni delle basi sono pari all'ampiezza delle classi e i cui volumi sono uguali o proporzionali alle frequenze.

»» Tabella a doppia entrata

Tabella statistica in cui figurano le frequenze assolute o relative o le probabilità riguardanti le diverse combinazioni di modalità o classi di modalità di due caratteri X e Y, desumibili da una distribuzione doppia.

»» Tabella statistica

Tabella derivante dalla classificazione dei dati statistici, e tale che ad ogni modalità o classe di modalità di uno o più caratteri corrisponde la rispettiva frequenza assoluta o relativa, o la rispettiva probabilità.

Una tabella statistica può essere semplice o multipla a seconda che figurino le modalità relative ad un solo carattere o a più caratteri.

»» Unità campionaria

Una delle n unità che fanno parte del campione. Da non confondere con l'unità di campionamento.

»» Unità di campionamento

Unità di insieme da sottoporre a campionamento.

»» Unità statistica

Componente elementare della popolazione su cui si acquisiscono informazioni. Le unità statistiche si distinguono in:

- *unità semplici* come una singola persona, una singola abitazione ecc.;
- *unità composte* se sono insiemi di unità semplici simili considerate anche a prescindere dall'unità composta;
- *unità complesse* se sono insiemi di unità semplici diverse considerate, però, nella loro globalità.

»» Valore centrale

Indice di posizione ottenuto facendo la semisomma tra valore minimo e massimo di una distribuzione di dati.

»» Valore centrale delle classi

Semisomma dei limiti superiore e inferiore di una classe di modalità.

I valori centrali sono utilizzati ai fini della determinazione di medie e di indici di variabilità in luogo delle modalità x_i quando i dati sono raggruppati in classi.

»» Variabile

In termini generali, si dice variabile un fenomeno (o una quantità) che assume valori diversi nel tempo, nello spazio o sotto circostanze differenti.

»» Variabile casuale continua

Una variabile casuale è continua se i valori che può assumere sono tutti quelli di un intervallo reale. Ad essa è associata una funzione di densità.

»» Variabile casuale discreta

Una variabile casuale è discreta se i valori che essa assume sono in corrispondenza di un insieme numerabile.

»» Variabilità

Attitudine di un carattere X ad assumere diverse modalità. Si distinguono due aspetti della variabilità:

- 1) la dispersione che caratterizza il maggiore o minore addensamento delle osservazioni intorno ad una media;
- 2) la disuguaglianza che evidenzia la diversità delle varie osservazioni tra loro. È misurata mediante gli *indici di variabilità*.

»» Varianza

Indice di variabilità molto utilizzato e indicato con σ^2 . Esso è definito come la media aritmetica del quadrato degli scarti dalla media aritmetica.

Rappresenta il quadrato dello scarto quadratico medio.